

PS Claim 61; Page 651; 1481pp; English.

XX The present sequence is a H. pylori cytoplasmic protein. The protein may  
CC be used in a vaccine to prevent or treat H. pylori infection or to  
CC identify H. pylori polypeptide binding compounds, useful as potential H.  
CC pylori life cycle activators or inhibitors. The genomic sequence of H.  
CC pylori (ATCC 55879) was determined from overlapping contigs generated by  
CC mechanically shearing the bacterial DNA. The sequences were analysed for  
CC ORF of at least 180 nucleotides, and the predicted coding regions defined  
CC by computer evaluation. To identify likely H. pylori antigens for vaccine  
CC development, the amino acid sequences predicted from various ORF were  
CC analysed for significant homology to other known or exported membrane  
CC proteins. Having identified and determined the sequences of interest,  
CC particular regions can be isolated from H. pylori by PCR amplification  
CC for recombinant polypeptide production, e.g. in E. coli hosts

XX Sequence 253 AA:

Query March 99.3%; Score 1270; DB 2; Length 253;  
Best Local Similarity 99.2%; Pred. No. 1.4e-117;  
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAKXDRLEFLKQLESSDLDLPEVLVFGKDGKRNKLTSSIRYKRGDPAKVAER 60  
DB 1 MAAKXDRLEFLKQLESSDLDLPEVLVFGKDGKRNKLTSSIRYKRGDPAKVAER 60  
QY 61 IAEELQYYSNSFASPIKSGCVLYKELCDVCDKLVNNYKKTETTLIEQNMLSKILERS 120  
DB 61 IAEELQYYSNSFASPIKSGCVLYKELCDVCDKLVNNYKKTETTLIEQNMLSKILERS 120  
QY 121 LEEMDEBEVKEMCDRLSIKNTDNLNRQALSAAATLTLPKMGKPSYOLAVIVANAATKTL 180  
DB 121 LEEMDEBEVKEMCDRLSIKNTDNLNRQALSAAATLTLPKMGKPSYOLAVIVANAATKTL 180  
QY 181 GRGLSLAGNQLTTLTSLPLTGPVGMITGWTADIDAGPARYVTIPACIVATLTKTQQ 240  
DB 181 GRGLSLAGNQLTTLTSLPLTGPVGMITGWTADIDAGPARYVTIPACIVATLTKTQQ 240

QY 241 ANGDKKSLQIESI 253  
DB 241 ANEDKKSQIESV 253

RESULT 4  
ID AAW24673 standard; protein: 253 AA.

XX AAW24673  
XX 12-AUG-1997 (first entry)

DE H. pylori cytoplasmic protein, 4095342.aa.  
XX Transmembrane; cytoplasmic; cell envelope; flagella; transport; secreted;  
XX periplasmic; chronic gastritis; duodenal ulcer disease; activator;  
XX inhibitor; bacterial life cycle; vaccine; immune; detection; antiserum;  
XX inhibition.

OS Helicobacter pylori.

XX MO9719098-A1.

XX 29-MAY-1997.

XX 15-NOV-1996; 96WO-US018542.

XX 17-NOV-1995; 95US-00561469.

XX (ASTR ) ASTRA AB.

XX Smith DH;

XX MPI, 1997-298052/27.

DR N-PSDB; AAT77491.

XX Helicobacter pylori nucleic acid sequences and related proteins - used  
PT for diagnostics and therapeutic.

XX Claim 18; Page 184; 235pp; English.

XX This sequence represents an H. pylori cytoplasmic protein. Helicobacter  
CC pylori has been strongly linked to chronic gastritis and duodenal ulcer  
CC disease. The nucleic acid sequences of the invention are used to evaluate  
CC compounds, especially activators or inhibitors of bacterial life cycle,  
CC for the ability to bind an H. pylori nucleic acid sequence. The nucleic  
CC acid sequences and corresponding proteins, are also useful for  
CC generating vaccines for immunising subjects against H. pylori or for use  
CC in detecting the presence of Helicobacter species in a sample. Antisense  
CC nucleic acid sequences of these sequences are used to inhibit expression  
CC of a gene from Helicobacter species. H. pylori whole genomic DNA was  
CC isolated and nebulized to a median size of 2000 bp. Purified DNA  
CC fragments were blunt-ended and ligated to unique BstXI-linker adapters in  
CC 100-1000 fold molar excess. These linkers are complementary to the BstXI-  
CC cut pMP vectors, while the overhang is not self-complementary. Therefore  
CC the linkers will not concatamerise nor will the cut vector re-ligate  
CC itself easily. The linker-adaptor inserts were ligated to each of the 20  
CC pMP vectors to construct a series of shotgun subclone libraries. The  
CC purified DNA samples were then sequenced. Note: The ORF/protein reference  
CC number for this sequence was obtained from the related specification,  
CC WO9640893

XX Sequence 253 AA:

Query March 99.3%; Score 1270; DB 2; Length 253;  
Best Local Similarity 99.2%; Pred. No. 1.4e-117;  
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAKXDRLEFLKQLESSDLDLPEVLVFGKDGKRNKLTSSIRYKRGDPAKVAER 60  
DB 1 MAAKXDRLEFLKQLESSDLDLPEVLVFGKDGKRNKLTSSIRYKRGDPAKVAER 60  
QY 61 IAEELQYYSNSFASPIKSGCVLYKELCDVCDKLVNNYKKTETTLIEQNMLSKILERS 120  
DB 61 IAEELQYYSNSFASPIKSGCVLYKELCDVCDKLVNNYKKTETTLIEQNMLSKILERS 120  
QY 121 LEEMDEBEVKEMCDRLSIKNTDNLNRQALSAAATLTLPKMGKPSYOLAVIVANAATKTL 180  
DB 121 LEEMDEBEVKEMCDRLSIKNTDNLNRQALSAAATLTLPKMGKPSYOLAVIVANAATKTL 180  
QY 181 GRGLSLAGNQLTTLTSLPLTGPVGMITGWTADIDAGPARYVTIPACIVATLTKTQQ 240  
DB 181 GRGLSLAGNQLTTLTSLPLTGPVGMITGWTADIDAGPARYVTIPACIVATLTKTQQ 240

QY 241 ANGDKKSLQIESI 253  
DB 241 ANEDKKSQIESV 253

RESULT 5  
ID AAW20866 standard; protein: 256 AA.

XX AAW20866;

XX 18-JUL-1997 (first entry)

XX H. pylori cytoplasmic protein, 129620305orf30.

XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;  
XX identification; binding compound; bacterium; life cycle; activator;  
XX bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;  
XX diagnosis.

XX Helicobacter pylori.

XX Smith DH;

XX WO9640893-A1.

DB 161 GAGLSLAGNOVLTFTLPGVOM::ITGVMTADIDAGPAVVTIPACIVATLRLKTOQ 240  
QY 241 ANGDKKSLOIISI 253  
DB 241 ANGDKKSLOIISI 253

## RESULT 2

US-09-732-091-44  
Sequence 44, Application US/09732091  
Patent No. US2002010736A1  
GENERAL INFORMATION:  
APPLICANT: Tiao, Jing-Fui  
APPLICANT: Walker, Richard I.  
APPLICANT: Jackson, W. James  
TITLE OF INVENTION: Helicobacter proteins, gene sequences and uses  
TITLE OF INVENTION: thereof  
FILE REFERENCE: 7969-088  
CURRENT APPLICATION NUMBER: US/09/732.091  
CURRENT FILING DATE: 2000-12-07  
NUMBER OF SEQ ID NOS: 44  
SEQUENCE: Patent in Ver. 2.1  
SEQ ID NO 44  
LENGTH: 265  
TYPE: PAT  
ORGANISM: Helicobacter sp.  
US-09-732-091-44

Query Match 100.0%; Score 1279; DB 9; Length 265;  
Best Local Similarity 100.0%; Freq. No. 5.1e-115;  
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYTJSDLEFLKQESSDLDLEFLVYFGKDEKRNKLTSSISFYKRGDDYATYAR 60  
DB 13 MAYTJSDLEFLKQESSDLDLEFLVYFGKDEKRNKLTSSISFYKRGDDYATYAR 72  
QY 61 IAELOVYGSNSPASFYKGEVLYKELCDVCDLAVNNKKTETTL2EONMLSKILERS 120  
DB 73 IAELOVYGSNSPASFYKGEVLYKELCDVCDLAVNNKKTETTL2EONMLSKILERS 132  
QY 121 LEBMDSEVEMCBLSIKNTDNLNRQALSATLTLPRKGGFSYGLAVYVANAVALTIL 180  
DB 133 LEBMDSEVEMCBLSIKNTDNLNRQALSATLTLPRKGGFSYGLAVYVANAVALTIL 192  
QY 191 GAGLSLAGNOVLTFTLPGVOM::ITGVMTADIDAGPAVVTIPACIVATLRLKTOQ 240  
DB 193 GAGLSLAGNOVLTFTLPGVOM::ITGVMTADIDAGPAVVTIPACIVATLRLKTOQ 252  
QY 241 ANGDKKSLOIISI 253  
DB 253 ANGDKKSLOIISI 253

## RESULT 3

US-10-335-977-9162  
Sequence 9162, Application US/10335977  
Publication No. US20040052799A1  
GENERAL INFORMATION:  
APPLICANT: DOUGLAS SMITH et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS

## NUMBER OF SEQUENCES: 10031

## CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE &amp; COCKFIELD

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM, ISO9660

COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: Windows NT 4.0  
SOFTWARE: UNIX  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/335.977  
FILING DATE: 30-Dec-2002  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/993.002  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragoras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: GTN-018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 9162:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: mlec feature  
LOCATION: (8) LOCATION 1...253  
SEQUENCE DESCRIPTION: SEQ ID NO: 9162:  
US-10-335-977-9162

Query Match 99.3%; Score 1270; DB 12; Length 253;  
Best Local Similarity 99.2%; Freq. No. 3.6e-114;  
Matches 251; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAYTJSDLEFLKQESSDLDLEFLVYFGKDEKRNKLTSSISFYKRGDDYATYAR 60  
DB 13 MAYTJSDLEFLKQESSDLDLEFLVYFGKDEKRNKLTSSISFYKRGDDYATYAR 72  
QY 61 IAELOVYGSNSPASFYKGEVLYKELCDVCDLAVNNKKTETTL2EONMLSKILERS 120  
DB 61 IAELOVYGSNSPASFYKGEVLYKELCDVCDLAVNNKKTETTL2EONMLSKILERS 120  
QY 121 LEBMDSEVEMCBLSIKNTDNLNRQALSATLTLPRKGGFSYGLAVYVANAVALTIL 180  
DB 121 LEBMDSEVEMCBLSIKNTDNLNRQALSATLTLPRKGGFSYGLAVYVANAVALTIL 180  
QY 191 GAGLSLAGNOVLTFTLPGVOM::ITGVMTADIDAGPAVVTIPACIVATLRLKTOQ 240  
DB 193 GAGLSLAGNOVLTFTLPGVOM::ITGVMTADIDAGPAVVTIPACIVATLRLKTOQ 240  
QY 241 ANGDKKSLOIISI 253  
DB 241 ANGDKKSLOIISI 253

## RESULT 4

US-10-335-977-9163  
Sequence 9163, Application US/10335977  
Publication No. US20040052799A1  
GENERAL INFORMATION:  
APPLICANT: DOUGLAS SMITH et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS

## NUMBER OF SEQUENCES: 10031

## CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE &amp; COCKFIELD

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

CC Helicobacter sp. HP30 protein  
XX Sequence 253 AA;

Query Match 100.0%; Score 1279; DB 5; Length 253;  
Beat Local Similarity 100.0%; Pred. No. 1.9e-118;  
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYKRDLEFLKQLSSDLDLFEVLVFGKGGKRNKLTSSIEYKRGDDYAKYAE 60  
DB 1 MAYKRDLEFLKQLSSDLDLFEVLVFGKGGKRNKLTSSIEYKRGDDYAKYAE 60  
QY 61 IAEELQYGSNSPASPFGEGVLYKEILCDVCDKLVNKKETTLIEQNMLSKILERS 120  
DB 61 IAEELQYGSNSPASPFGEGVLYKEILCDVCDKLVNKKETTLIEQNMLSKILERS 120  
QY 121 LEEEMDEEYKMCDELSTKNTDNLNROALSAATLTLFPMGCFKSYQLAVIVANNAVKTL 180  
DB 121 LEEEMDEEYKMCDELSTKNTDNLNROALSAATLTLFPMGCFKSYQLAVIVANNAVKTL 180  
QY 181 GRGLSLAGNOVLTRTLSTLPGVGMIIIGVWTAIDAGPAVRVTIPACIVATRLKTOO 240  
DB 181 GRGLSLAGNOVLTRTLSTLPGVGMIIIGVWTAIDAGPAVRVTIPACIVATRLKTOO 240  
QY 241 ANGDKKSQIEST 253  
DB 241 ANGDKKSQIEST 253

## RESULT 2

AA26878 standard; protein; 265 AA.

AC AAE26878;

DT 13-DEC-2002 (first entry)

XX Helicobacter sp. PGB/HP30 protein.

XX HP30; HP56; immune response; therapy; Helicobacter infection; vaccine;  
KM type B gastritis; antiinflammatory; adenocarcinoma; defense mechanism;  
KM low grade B cell lymphoma; virulence; antibacterial; gastric cancer;  
KM immunostimulant; cytotoxic; peptic ulcer.

OS Helicobacter sp.

XX Key Location/Qualifiers  
FH Misc-difference 12  
PT /note= "Encoded by TTC"

XX MO20251237-A2.

PD 04-JUL-2002.

XX 07-DEC-2001, 2001MO-US048392.

XX 07-DEC-2000, 2000US-00732091.

XX (ANTE-) ANTEK BIOLOGICS INC.

XX Tian U, Walker R, Jackson WJ;

XX MPI, 2002-666854/71.

XX N-PSDB; AAD44535.

XX Novel Helicobacter proteins, HP30 and HP56, and nucleic acid encoding  
PT the proteins, useful as vaccines for raising immune response in animals.

XX Disclosure: Page 122-123; 127pp; English.

XX The invention relates to Helicobacter HP30 or HP56 polypeptide and  
CC peptides derived from them. The invention is useful for producing an  
CC immune response. It is useful for preventing, treating or ameliorating a

CC disorder or disease associated with infection of an animal with  
CC Helicobacter. Pharmaceutical composition and vaccines comprising the  
CC sequence of the invention is useful for treating type B gastritis,  
CC peptic ulcers, gastric cancers such as adenocarcinoma, and low grade B  
CC cell lymphoma. The invention is useful as reagents for clinical or  
CC medical diagnosis of Helicobacter infections and for scientific research  
CC on the properties of pathogenicity, virulence and infectivity of  
CC Helicobacter, as well as host defense mechanisms. The present sequence is  
CC Helicobacter sp. PGB/HP30 protein

SO Sequence 265 AA;

Query Match 100.0%; Score 1279; DB 5; Length 265;  
Beat Local Similarity 100.0%; Pred. No. 1.9e-118;  
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYKRDLEFLKQLSSDLDLFEVLVFGKGGKRNKLTSSIEYKRGDDYAKYAE 60  
DB 1 MAYKRDLEFLKQLSSDLDLFEVLVFGKGGKRNKLTSSIEYKRGDDYAKYAE 72  
QY 61 IAEELQYGSNSPASPFGEGVLYKEILCDVCDKLVNKKETTLIEQNMLSKILERS 120  
DB 73 IAEELQYGSNSPASPFGEGVLYKEILCDVCDKLVNKKETTLIEQNMLSKILERS 132  
QY 121 LEEEMDEEYKMCDELSTKNTDNLNROALSAATLTLFPMGCFKSYQLAVIVANNAVKTL 180  
DB 133 LEEEMDEEYKMCDELSTKNTDNLNROALSAATLTLFPMGCFKSYQLAVIVANNAVKTL 192  
QY 181 GRGLSLAGNOVLTRTLSTLPGVGMIIIGVWTAIDAGPAVRVTIPACIVATRLKTOO 240  
DB 193 GRGLSLAGNOVLTRTLSTLPGVGMIIIGVWTAIDAGPAVRVTIPACIVATRLKTOO 252  
QY 241 ANGDKKSQIEST 253  
DB 253 ANGDKKSQIEST 265

## RESULT 3

AA20486 standard; protein; 253 AA.

AC AAM20486;

DT 29-JUL-1997 (first entry)

XX H. pylori cytoplasmic protein, 4095342.aa.

XX Cytoplasmic; vaccine; prevention; treatment; infection; identification;  
KM binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
KM duodenal ulcer disease; chronic gastritis; diagnosis; envelope.

XX Helicobacter pylori.

XX MO9640893-A1.

PD 19-DEC-1996.

XX 06-JUN-1996; 96MO-US009122.

XX 07-JUN-1995; 95US-00487032.

XX 01-APR-1996; 96US-00630405.

XX (ASTR) ASTRA AB.

XX Smith D, Berglund OT, Wellgaard BJ,

XX MPI, 1997-052306/05.

XX N-PSDB; AAT67811.

XX Helicobacter pylori nucleic acid sequences and related polypeptide(s)  
PT useful for vaccines to treat or prevent H. pylori infection, and to  
PT detect Helicobacter.

1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 13 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 120  
 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 52  
 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 72  
 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 92  
 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 112  
 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 132  
 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 152  
 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 172  
 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 192  
 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 212  
 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 232  
 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTAACGCTTAACGCAAGCA 720  
 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 252  
 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 265

RESULT 3  
 ID AAM20486 standard; protein; 253 AA.

AC AAM20486  
 XX  
 XX 29-JUL-1997 (first entry)  
 XX  
 XX H. pylori cytoplasmic protein, 4095342.aa.  
 XX  
 XX Cytoplasmic vaccine; prevention; treatment; infection; identification;  
 XX binding compound; bacterium; life cycle; activator; bacteria; inhibitor;  
 XX duodenal ulcer disease; chronic gastritis; diagnosis; envelope.  
 XX  
 XX Helicobacter pylori.  
 XX  
 XX MO640893-A1.  
 XX  
 XX 19-DEC-1996.  
 XX  
 XX 06-JUN-1996; 96MO-UB009122.  
 XX  
 XX PF

XX  
 PR 07-JUN-1995; 95US-00487032.  
 PR 01-APR-1996; 96US-00630405.  
 XX  
 XX (ASTR) ASTRA AB.  
 XX  
 XX Smith D, Berglindh OT, Mellgard BL;  
 XX WPI; 1997-052306/05.  
 DR N-PSDB; AAT67811.  
 XX  
 PT Helicobacter pylori nucleic acid sequences and related polypeptide(s) -  
 PT useful for vaccines to treat or prevent H. pylori infection, and to  
 XX detect Helicobacter.  
 PS  
 PS Claim 61: Page 651; 1481pp; English.  
 XX  
 XX The present sequence is a H. pylori cytoplasmic protein. The protein may  
 CC be used in a vaccine to prevent or treat H. pylori infection or to  
 CC identify H. pylori polypeptide binding compounds, useful as potential H.  
 CC pylori life cycle activators or inhibitors. The genomic sequence of H.  
 CC pylori (ATCC 55679) was determined from overlapping contigs generated by  
 CC ORF of at least 180 nucleotides, and the predicted coding regions defined  
 CC by computer evaluation. To identify likely H. pylori antigens for vaccine  
 CC development, the amino acid sequences predicted from various ORF were  
 CC analysed for significant homology to other known or exported membrane  
 CC proteins. Having identified and determined the sequences of interest,  
 CC particular regions can be isolated from H. pylori by PCR amplification  
 CC for recombinant polypeptide production, e.g. in E. coli hosts  
 XX  
 XX Sequence 253 AA:  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 6,946-134 Length: 253  
 XX Score: 1270.00 Matches: 253  
 XX Percent Similarity: 99.60% Conservative: 1  
 XX Best Local Similarity: 99.21% Mismatches: 0  
 XX Query Match: 94.14% Indels: 0  
 XX DB: 2 Gaps: 0  
 XX  
 XX US-09-732-091-3 (1-759) x AAM20486 (1-253)  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCTTTCTAAATCTTAAGAAAGT 360  
 DB 113 LysLysThrGlnThrThrLeuIleGlnGlnAspMetLeuSerTyrIleLeuGlnLysSer 120  
 XX  
 XX 361 TTGCAAGAAATGATGATTAAGCAAGCAATGCGCAATTCGCAATTAATTCGCAAGAA 420  
 DB 133 LeuGlnGlnLeuAspArgGlnGlnValTyrGlnMetCysAspGlnLeuSerIleLysAsn 140  
 XX  
 XX 421 ACGCAATTTAAACACACAGCTTAAGCGCGCAATTAACGCTTTAAATGCGAA 480  
 DB 153 ThrAspLeuLeuAspArgGlnAlaLeuSerLeuIleThrLeuThrLeuLeuLeuLeuLeu 150  
 XX  
 XX 481 GATTAACTTATCAATTAAGCTGCTATTTGCAATGCGTGCAGAAACCAATTCCTA 540  
 DB 173 GlyPheLysSerTyrGlnLeuAlaValIleValAlaAlaAlaAlaAlaAlaAlaAla 160  
 XX  
 XX 541 GCGCGCTGTTACGCTTGGCGGCAATGAGTGTACACAACTGAGCTTTTAAACA 600  
 DB 193 GlyArgGlyGlnSerLeuAlaGlnValLeuThrAspThrLeuSerPheLeuThr 180  
 XX  
 XX 601 GCTCCTGTTGAGTGAATCATTAACGCGGCAATGAGCAAGCAATTCGCAAGCGCGCT 660  
 DB 213 GlyProGlyGlyThrIleThrIleThrGlyValTyrThrAlaIleAspThrAlaGlyProAla 170  
 XX  
 XX 661 TATAGCTTAACGTAACGCGGCAATGAGTGTGCTTACGCTTACGCTTAACGCAAGCA 720  
 DB 233 TyrArgValThrIleProAlaCysIleValAlaThrLeuArgLeuLysThrGlnGln 210  
 XX  
 XX 721 GCGAATGAGATTAAGAGTGTGCAATTAACGCAATTCATT 759  
 DB 253 AlaAsnGlyAspLysSerLeuGlnIleGlnSerIle 240  
 XX  
 XX 1 ATGGCATTAACAAATATGATAGAGCTTGCAATTTTAAAGCAATTCGAAATGATATTA 60  
 DB 1 MetAlaTtlytlytAspArgAspLeuGlnPheLeuGlnLeuGlnSerSerAspLeu 20  
 XX  
 XX 61 TTGGATTGTTGAGGCTGCTTTGTTTGTAAAGACCGCCGCAAAAACACACAGCAAGAAA 120  
 DB 33 LeuAspPhePheGluValLeuValPheGlyAspGlyGlyValArgGlyGlyValGly 40  
 XX  
 XX 121 CTGACCAAGCTCATAGATTAACAAAGCAATGCGCAATTAACGTAATTCGCAAGAA 180  
 DB 53 LeuThrSerSerLeuGlnTyrLeuAspArgAspLeuAspArgAspLeuAspArgAsp 60  
 XX  
 XX 181 ATCCGTAAGAGTTGCAATTAACGTAAGCAATGCGCAATTCGCAATTAACGTAAGCA 240  
 DB 73 IleAlaGlnGlnLeuGlnTyrTyrGlySerAspSerPheIleSerPheIleGlyGlu 80  
 XX  
 XX 241 GGAGCTTATACAAAGCAATTTATGCAATGCTGCAATTAATTAAGTCAATTAACAC 300  
 DB 93 GlyValLeuTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyrLeuGlnTyr 100  
 XX  
 XX 301 AAGAAACCTGAACGCAATTAATTAACAAAGCAATGCT